AFRL-RH-BR-TR-2009-0031

24/7 OPERATIONAL EFFECTIVENESS TOOLSET: MISHAP INVESTIGATION INTERFACE

Douglas R. Eddy
James C. Miller
Cory Welch
Richard Smith
Samuel L. Moise

NTI, Incorporated
1 ½ South Central Ave
Fairborn, OH 45324

October 2008

Interim Report for Mar 2006 – Aug 2008

Air Force Research Laboratory
711 Human Performance Wing
Human Effectiveness Directorate
Biosciences and Protection Division
Biobehavioral Performance Branch
Brooks City-Base, TX 78235

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 311th Public Affairs Office at Brooks City Base, TX and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-BR-TR-2009-0031 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

_____//SIGNED//_____   _____//SIGNED//_____
SHARON K. GARCIA   MARK M. HOFFMAN
Work Unit Monitor   Deputy Division Chief
Biobehavioral Performance Branch   Biosciences and Protection Division
                                         Human Effectiveness Directorate
                                         711 Human Performance Wing
                                         Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
15-10-2008	Interim Technical Report	30-03-2006 to 29-08-2008

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
24/7 Operational Effectiveness Toolset: Mishap Investigation Interface	FA8650-06-C-6606
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
Douglas R. Eddy , James C. Miller, Cory Welch, Richard Smith, and Samuel L. Moise	5020
	5e. TASK NUMBER
	P9
	5f. WORK UNIT NUMBER
	09

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
NTI, Inc. 1 ½ S. Central Ave. Fairborn, OH 65324	

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
	AFRL-RH-BR-TR-2009-0031

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report describes the Mishap Investigation Interface of the 24/7 Operational Toolset, a World Wide Web-based fatigue analysis software product. The toolset was based upon the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE™; Hursh et al., 2004). The SAFTE™ model predicts cognitive performance level based upon sleep, circadian rhythm, and sleep inertia. This specific interface of the toolset was designed to aid in the determination of whether aviation and ground mishaps may have causes rooted in or exacerbated by mental fatigue. The interface design approach was iterative, involving several meetings among subject matter experts (SMEs), interface software designers and evaluators. The first meeting was for the purpose of requirements analysis, in which the designers elicited task information from the SMEs. The second meeting included a walk-through of storyboarded and preliminary software, in which the SMEs provided feedback to the designers and evaluators. The final meeting was for the purpose of an "inspection evaluation" of the interface by the SMEs and evaluators. This interface was based upon task analyses of mishap investigators who also served as our SMEs. The requirements analysis revealed how fatigue related information is entered into the report of an Air Force Safety Investigation Board. A task analysis revealed how safety investigators accomplish their investigation. The report begins with a general review of mishap investigation and proceeds to the roles played by our SMEs in determining the effects of fatigue. The walk-through and inspection evaluation processes indicated that most of the requirements of potential users were met reasonably well and that potential users were able to operate the interface reasonably easily. The report ends with recommendations and suggestions for incorporating a fatigue-assessment decision aid into the mishap investigation process.

15. SUBJECT TERMS

Mishap investigation, Fatigue, Cognitive performance, Task analysis, Sleep, Circadian rhythm, Interface design

16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Dr. Sharon Garcia
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	SAR	55	19b. TELEPHONE NUMBER *(include area code)*

i

This page left intentionally blank

PREFACE

The technical monitors for this work were Dr. James C. Miller and Scott Chaiken and the program managers were 1Lt Andrew Workman and 1Lt Andrea Pinchak, all of the Biosciences and Protection Division, Air Force Research Laboratory at Brooks City-Base, Texas. 1Lt Pinchak coordinated several meetings with the mishap investigators at USAFSAM. The success of this endeavor was only made possible through the generous support of LtCol Andrew Woodrow and his colleagues at USAFSAM at Brooks City-Base. Their in-depth comments, insightful recommendation, and thoughtful suggestions were critical to designing a useful interface for mishap investigators. Early in the effort, Dr. William F. Storm assisted with the task analyses and Mr. Juan Mendez researched information on the internet, tested early versions of the software, and we thank them both for their contributions. Additional thanks go to Beth Barker for handling financial relationships with the government and our subcontractors.

TABLE OF CONTENTS

FIGURES

TABLES

SUMMARY

This report describes the Mishap Investigation Interface of the 24/7 Operational Toolset, a World Wide Web-based fatigue analysis software product. The toolset was based upon the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE™; Hursh et al., 2004). The SAFTE™ model predicts cognitive performance level based upon sleep, circadian rhythm, and sleep inertia. This specific interface of the toolset was designed to aid in the determination of whether aviation and ground mishaps may have causes rooted in or exacerbated by mental fatigue. The interface design approach was iterative, involving several meetings among subject matter experts (SMEs), interface software designers and evaluators. The first meeting was for the purpose of requirements analysis, in which the designers elicited task information from the SMEs. The second meeting included a walk-through of storyboarded and preliminary software, in which the SMEs provided feedback to the designers and evaluators. The final meeting was for the purpose of an "inspection evaluation" of the interface by the SMEs and evaluators. This interface was based upon task analyses of mishap investigators who also served as our SMEs. The requirements analysis revealed how fatigue related information is entered into the report of an Air Force Safety Investigation Board. A task analysis revealed how safety investigators accomplish their investigation. The report begins with a general review of mishap investigation and proceeds to the roles played by our SMEs in determining the effects of fatigue. The walk-through and inspection evaluation processes indicated that most of the requirements of potential users were met reasonably well and that potential users were able to operate the interface reasonably easily. The report ends with recommendations and suggestions for incorporating a fatigue-assessment decision aid into the mishap investigation process.

24/7 Operational Effectiveness Toolset:
Mishap Investigation Interface

INTRODUCTION

Department of Defense (DOD) aviation safety has improved significantly over the years. Between 1975 and 1995 for example, the annual number of Class A mishaps decreased from 309 to 76, while the number of fatalities decreased from 285 to 85. During this period, Class A mishaps per 100,000 flying hours, referred to as the mishap rate, also decreased from about 4.3 to 1.5. The value of Class A losses remained fairly constant over the following years, ranging from a high of about $1.6 billion in 1993 to a low of $1.2 billion in 1994. Each of the services have taken steps to reduce aviation mishaps, such as tracking mishap investigation recommendations and disseminating safety information in manuals, newsletters, videos, and messages. During the mid 1990's the services implemented safety initiatives including risk management and human factor studies (Gebicke ME, 1996). Recent fatigue modeling developments have led to the development of new tools for estimating the effects of fatigue on human performance.

This report describes the Mishap Investigation Interface of the 24/7 Operational Toolset and the approach used to create it. The toolset was based upon the Sleep, Activity, Fatigue, and Task Effectiveness model (SAFTE™; Hursh et al., 2004) and on the Fatigue Avoidance Scheduling Tool (FAST™, Eddy & Hursh, 2001, 2006a, 2006b). The SAFTE™ model predicts cognitive performance level based upon sleep, circadian rhythm, and sleep inertia. Figuring out what causes an airplane to crash is no easy task and this specific interface of the toolset was designed to aid in the determination of whether aviation and ground mishaps may have causes rooted in or exacerbated by mental fatigue. This interface was based upon task analyses of how mishap investigators do their job. The requirements analysis revealed how fatigue related information is entered into the report of an Air Force Safety Investigation Board. The report begins with a general review of mishap investigation and proceeds to the roles played by our subject-matter-experts (SMEs) in determining the effects of fatigue. The walk-through and inspection evaluation processes indicated that most of the requirements of potential users were met reasonably well and that potential users were able to operate the interface to obtain a fatigue analysis. The report ends with recommendations and suggestions for incorporating a fatigue-assessment decision aid into the mishap investigation process. To understand the mishap investigation process more fully, the authors consulted additional documents not cited in this report. A short bibliography of these documents is included at the end of the reference section.

OVERALL ACCIDENT INVESTIGATION PROCESSES

The purpose of a safety investigation is to determine all factors; human, materiel, and environmental that directly or indirectly contributed to the mishap. This information can be used by aircrew, equipment operators, supervisors, commanders, staffs, and designers to eliminate the cause factors and prevent recurrence of a similar mishap. A safety board can

be composed of as few as five members (for a Class A flight mishap) or have many additional members (Figure 1).
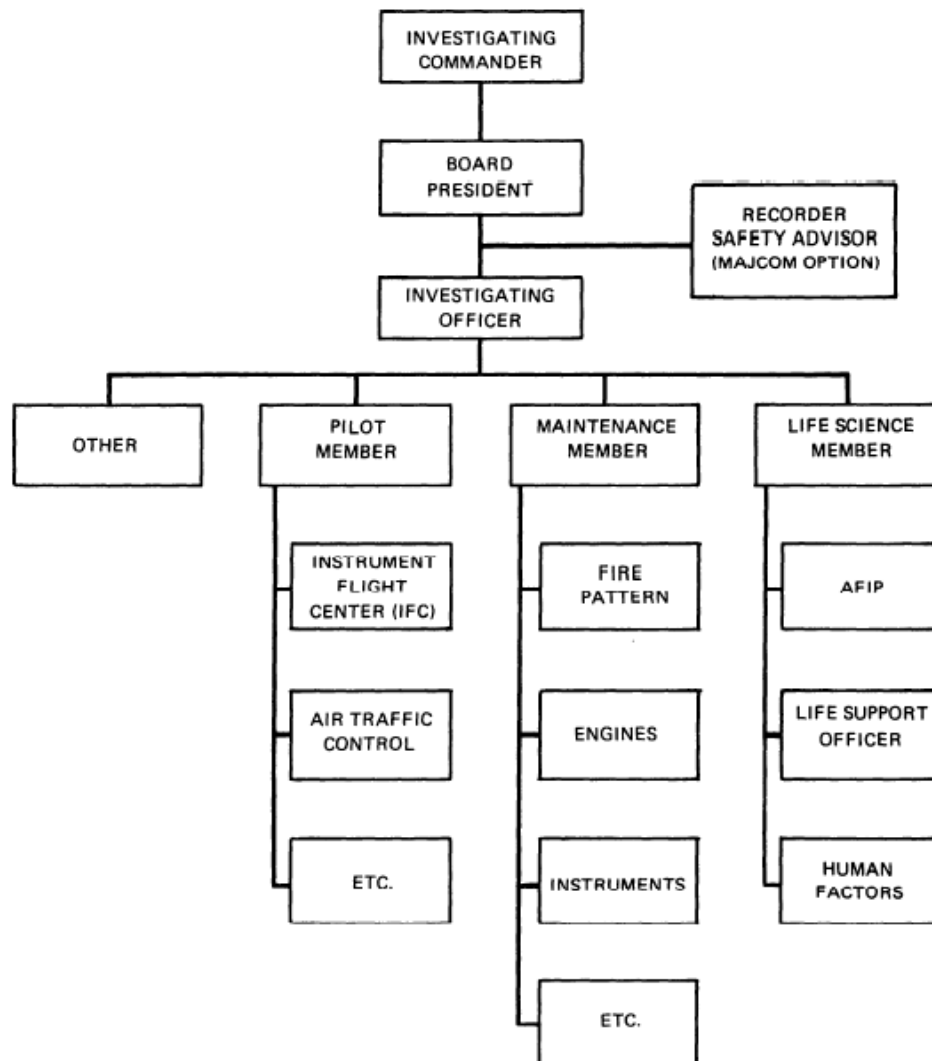
Figure 1. The organization of a formal Safety Investigation Board (AFP 127-1, Vol. 1 1987).

Different members of an investigative board have different qualifications to cover all the mishap's possible causes. Every board has an officer in charge (president). The president may be a lieutenant colonel or higher depending on the aircraft type (UAV vs. manned aircraft) and the occurrence of a fatality. The board will also have an investigating officer, a pilot member, a maintenance member, and a medical member and may have a safety center representative (non-voting). A safety center or other member may be an aerospace physiologist, psychologist, or human factors specialist. Each member contributes the results of their investigation to the report. Boards have invited specialized members in addition to the voting members. They provide their information on the people involved through the physician's portion of the report and the physician may edit their information.

The investigation process has six phases: preparation, notification, interim board action, arrival, investigation and analysis, and writing (USAF Guide to Mishap Investigation, 1987). This requirements assessment report briefly describes each phase, but focuses on the phases that might benefit from an automated fatigue assessment tool.

The preparation phase describes how an investigator prepares for the investigation task through training. The Interim Safety Board (ISB) phase involves preliminary investigation by units identified by a full-time safety staff in the first few hours after a mishap. The ISB makes sure perishable evidence is not lost, and that fluid samples, tower and RAPCON voice tapes, and aircraft and crew records are secured. They must do everything possible to help the follow-on Safety Investigation Board (SIB) get organized and started once they arrive. During the Arrival phase, the SIB plans its attack on the investigation to maximize the information acquired and minimize duplication of effort. A good investigation and report is the result of assigning SIB members to appropriate duties. The SIB spends most of its time in the Investigation and Analysis phase. Evidence is gathered, sorted, and evaluated. Each member carries out assigned duties acquiring and interpreting the information sought in their Arrival Phase planning.

During the writing phase SIB members sort out their notes, memos, and evidence and compile a clearly written and well documented report. They discard unneeded material and include meaningful evidence and analysis. The reasons for discarding apparent cause factors are clearly documented in the report.

MEDICAL AND HUMAN FACTORS PORTION OF MISHAP INVESTIGATION

As can be seen from Figure 1 the human factors analysis falls under the Life Science officer, usually a flight surgeon (FS). While the FS is tasked with performing the human factors analysis, they often need additional information usually beyond their expertise. Topic areas include supervisory and institutional concerns as well as more traditional human factors such as understanding the sequence of events, communication problems, peer influences, and access to adequate facilities and services. Supervisory issues include discipline enforcement, command and control, appropriate supervisory model behavior, and expressed pressure in tasking. Communication problems include those within the cockpit, between personalities, outside the cockpit, communications, and equipment failure. Peer influences include verbal comments, commonly held beliefs based on unspoken or unwritten learning, and perceptions of equipment concerns. Adequacy of access to quarters, nutrition, exercise, recreation and health care must also be examined. More directly, the facilities of an airfield or air traffic control services may have an impact. Although the FS must integrate these inputs into an overall human factors analysis, a life sciences officer with human factors expertise will likely write much of this section of the report (USAF Guide to Mishap Investigation, 1987; statements made by participants during TA.)

A life-support officer or physiological training officer (PTO) is normally appointed to the board when aircrew equipment, egress, or survival is involved in a mishap. This officer works directly with and for the FS. In addition to being "expert" in the life-support equipment, the PTO helps the FS understand the environment in which the aircrew

operates (i.e., operations, pressures, equipment limitations, attitudes, etc.). A PTO may have background or credentials in:

- Applied aerospace physiology
- Hyperbaric/hypobaric physiology training
- Disorientation and centrifuge training
- Aviation human factors
- Life-support course
- Parachutist and land or water survival course.

Drawing upon Reason's (1990) and Wiegmann and Shappell's (2003) concept of active failures and latent failures/conditions, the DoD developed a taxonomy to identify hazards and risks leading to a mishap called the DoD Human Factors Analysis and Classification System (HFACS). DOD-HFACS describes four main tiers of failures/conditions: 1) Acts, 2) Preconditions, 3) Supervision, and 4) Organizational Influences (Figure 2). The DoD HFACS model addresses human error from 3 perspectives: 1) cognitive viewpoint and human system interaction and integration, 2) human-to-human interaction, and 3) socio-cultural and organizational. The model focuses on preconditions and conditions of individuals' adverse physiological states and preconditions of personnel factors and self-imposed stress. With this approach, the investigator is attempting to determine if fatigue was a precondition to an unsafe act (e.g., a skill-based error, a judgment error, a perceptual error, or a violation). Once the investigator determines that fatigue may have been a precondition, he or she must then look at the supervisory and organizational levels to determine why. Some of the reasons might be inadequate manpower (an organizational influence) or crew scheduling (due to unsafe supervision).

The FS, sometimes with a consultant, generally handles the human performance area. If the need for specialized knowledge or experience is identified, HQ AF Safety Center (AFSC) can provide an expert that will ensure that a standardized analysis is conducted. When reporting to the board the expert must make clear the limits of current reliable knowledge in the area of concern and differentiate that from the necessary but subjective expressions of expert judgment. Evaluation techniques should include established practical methods, using standardized protocols, and avoid speculation. Statements of expert opinion are appropriate, but should be represented as such and not as fact (USAF Guide to Mishap Investigation, 1987).
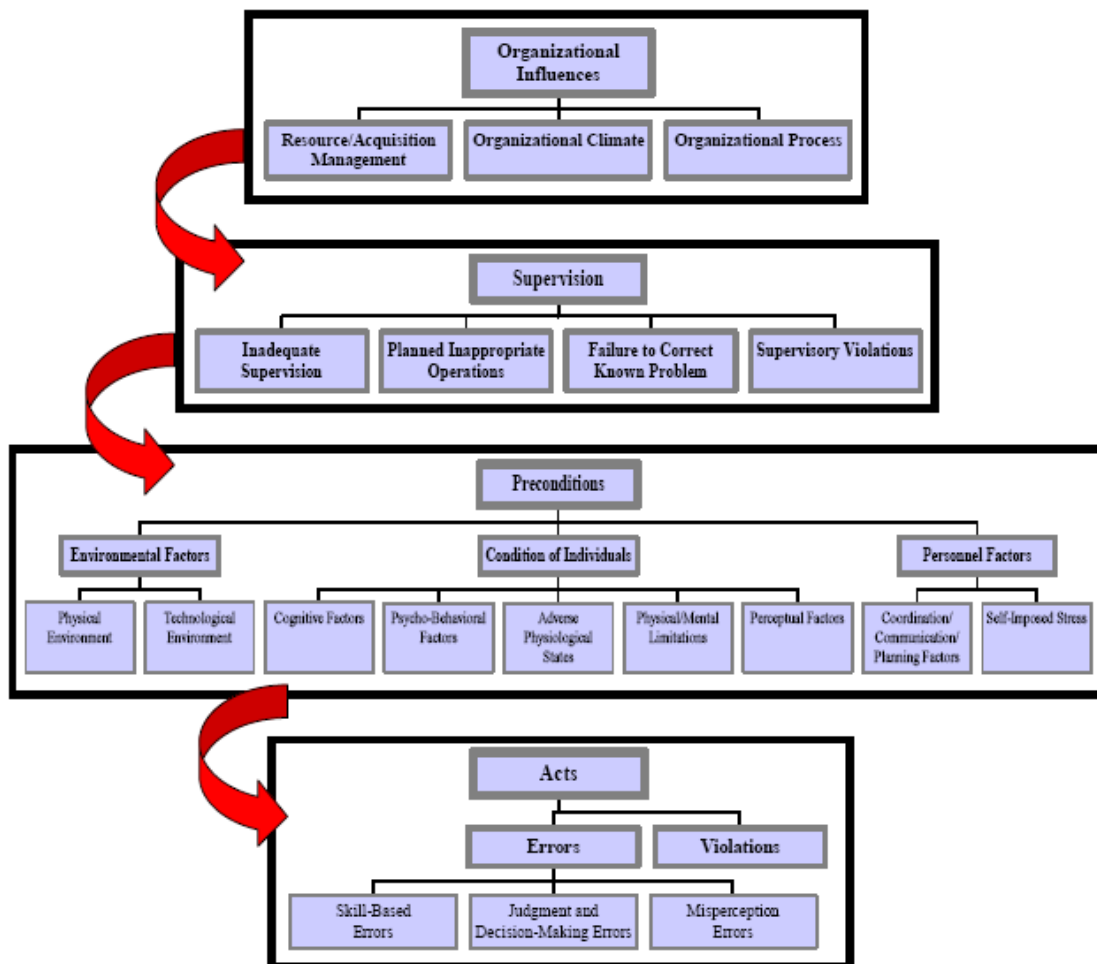
Figure 2. The DoD HFACS Model.

In the DoD HFACS model, fatigue falls under the categories labeled preconditions/conditions of individuals/adverse physiological states and preconditions/personnel factors/self-imposed stress. The specific fatigue nanocodes are: physical fatigue (overexertion), physiological/mental fatigue, circadian rhythm desynchrony, and inadequate rest. Definitions for each nanocode are included in the DoD HFACS document. In a memorandum dated 8 Sep 2004, AFRL/HEPF scientists forwarded revised definitions for the next version of the AFSC Human Factors Analysis and Classification System, which was under revision at that time. The following factors and definitions were included in a memo sent to AFSC and make identifying types of fatigue easier to discuss in the context of a mishap investigation. *Physical fatigue* degrades task performance and occurs when the individual's diminished physical capability is due to overexertion (time or relative load). Prolonged physical activity, or periods of brief, but relatively extreme, physical activity can severely tax a person's physical endurance or strength degrading their performance below their normal levels. *Acute mental fatigue* is a factor when the individual's diminished mental capability is due to a period of prolonged

wakefulness, usually more than 16 hours, occurring between two major sleep periods. *Cumulative mental fatigue* is considered a factor when the individual's diminished mental capability is due to disturbed or shortened major sleep periods occurring between two or more successive major waking, duty, or work periods. One major sleep period will not eliminate cumulative fatigue. The *circadian rhythm* itself degrades performance when the individual's normal, 24-hour, rhythmic biological cycle (circadian rhythm) conflicts with duty requirements. This is caused by one or more nights of work or rapid movement (faster than one time zone per day) across more than three time zones. These effects are referred to as "shift lag" and "jet lag," respectively. Continuous time spent in the new time zone will lead to acclimation, but more acclimation time is needed for each time zone crossed. Acclimation to night work may never occur. *Chronic mental fatigue* is a factor when the individual is frequently exposed to at least one month of multiple periods of prolonged wakefulness, excessive work hours, disturbed or shortened major sleep periods, unresolved conflicts, or prolonged frustration, and it degrades task performance. An individual must display, concurrently, four or more of the following symptoms: the desire to sleep, apathy, substantial impairment in short-term memory or concentration, muscle pain, multi-joint pain without swelling or redness, headaches of a new type, pattern or severity; sleep that does not refresh, or post-exertion malaise lasting for more than 24 hours. The symptoms must have persisted or recurred for at least one month. Chronic mental fatigue is not eliminated by any number of sleep periods without first removing the original cause.

Fatigue effects are pervasive (diminishing efficiency of mental processes from perception through exercise of judgment), ubiquitous (affect everything), and insidious (not known by the affected individual). Quantification of fatigue effects can be difficult, even for the expert. Having a model to predict performance degradation from sleep loss, circadian disruption, and sleep inertia will help to objectify the contribution of fatigue as a causal factor in a mishap.

COMPLETING THE MISHAP REPORT
The work of the SIB, as well as the ISB, is documented in a report using the AF Safety Automated System (AFSAS). AFSAS is a web-based software tool used to enter mishap reports, investigate existing mishaps, and to update and review existing mishap reports. The 9 March 2007 version of the system allows for the reporting of all mishaps: aviation, ground, or weapons[1]. The previous aviation user manual available at the time of this writing was dated November 17, 2003, and the system resides under the control of HQ AFSC, Kirtland AFB, NM, https://sas.kirtland.af.mil/. AFSAS provides the user with the following functions and features:
- Secure global access via the Internet with a different login name and password for each user.
- Main-menu access to different primary tasks such as creating, investigating, updating, and running mishap reports as well as administrating AFSAS accounts.

---

[1] Department of AF Memorandum from HQ AFSC/CD for MAJCOM/SEs and all AFSAS Users dated 13 February 2007, Greg Alston, Executive Director, https://sas.kirtland.af.mil/documents/AFSAS_Transition_Letter.pdf.

- Access tabs for displaying organized and interrelated user-interface screens.
- Data integrity insurance with a notifying message about any invalid or unspecified data.
- User-friendly data entry and retrieval with lists of options and search capabilities for most entries.
- Ad-hoc queries on the existing mishap reports with simple data entry forms and various query output reports.

A mishap report includes information on everything that may have contributed to it.
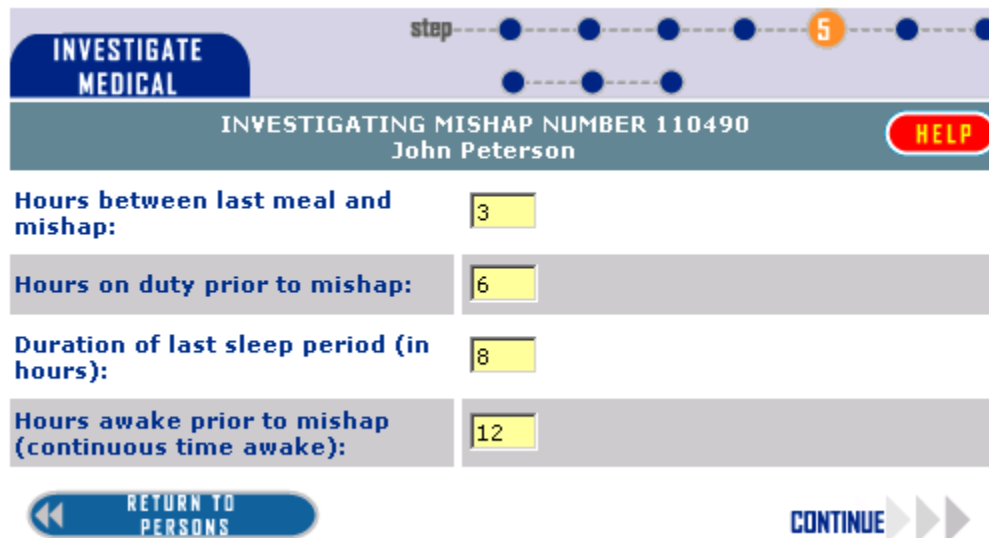- People
- Aircraft
- Conditions at the time
- Communications
- Documents
- Everything

However the DoD HFACS system goes further. Mishap reports should describe the unsafe act of the crew, maintenance, ATC, etc., the environmental, individual, and personnel preconditions, and latent failures at the supervisory and organizational levels as well. One of the key elements of HFACS is to identify latent failures temporally and physically distant from the actual mishap occurrence as these resident pathogens are likely to contribute to additional mishaps. For instance, while it is important to know that fatigue was contributory in a mishap, it is more important to know the crew was fatigued because the unit had inappropriate manning (organizational issue). However, if we don't identify fatigue, we will fail to follow the appropriate error trajectory and identify the latent supervisory and organizational failures.

As each board member enters information into the report, they follow a set of data input screens in AFSAS that prompt them for specific factual information that becomes a part of the archival mishap database. The human factors specialist investigates those personnel who had an active role in the mishap, were injured in the mishap, or whose actions or inactions initiated or sustained the mishap sequence. All of these people are included in the investigation and in AFSAS. The information on persons is divided into six different categories: demographic data, medical, equipment, survival/rescue, egress, and individual factors. All of the information in each of these categories is very important and is examined to determine causality. However, our focus is on the individual factors that may have affected the involved persons behavior, decision-making, accuracy, or speed of response. Questions regarding flying status, the length of assignment at the base, the number of days deployed over the past 365, the number of hours flown in the last 30, 60, and 90 day periods, and the level of cockpit management training are investigated.

Information concerning nutrition and sleep is entered in the medical portion of the AFSAS report. In the screen shown in Figure 3 the investigator enters, in hours, how long the person was on duty before the mishap, how long the person slept before the mishap, and how long the person stayed awake before the mishap occurred. The follow-on screen,

shown in Figure 4, asks for more details about sleep. Unfortunately, the information available from these screens is insufficient to enable a model to simulate the effects of sleep, or lack of it, on cognitive performance. Missing are the time and number of sleep periods, and the duration of each during the 72-hours prior to the mishap.
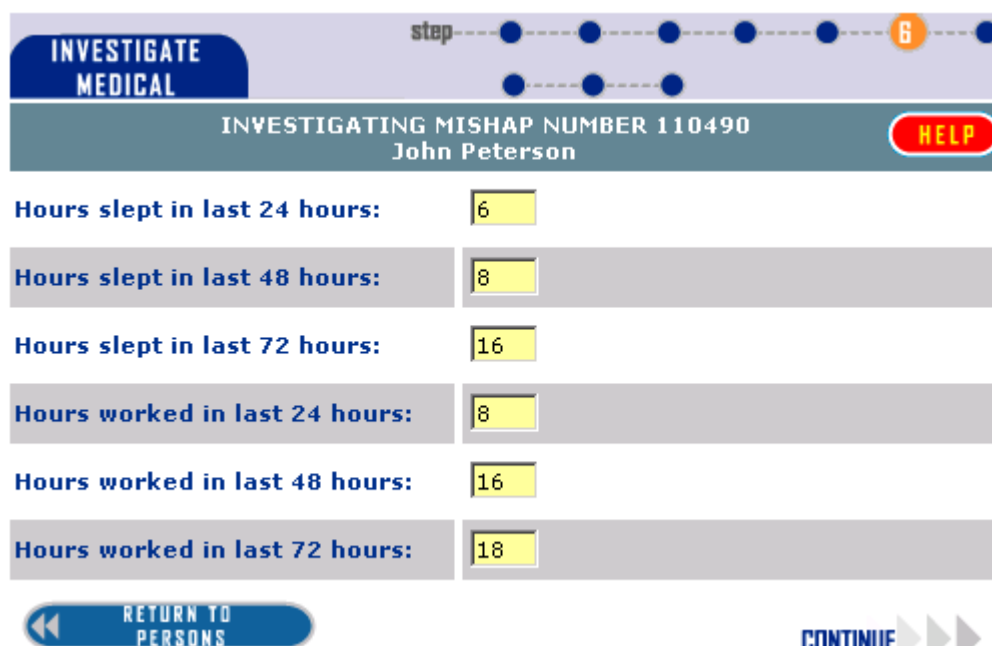
Figure 3. Current AFSAS entry screen for nutrition and sleep information.

Figure 4. Current AFSAS entry screen for "detailed" sleep information prior to the mishap.

A final screen, shown in Figure 5, asks the investigator to give an opinion on whether or not fatigue may have been a factor in the mishap. The instructions for completing this field only specify long-term fatigue as a possible factor. As long-term fatigue is not defined in the manual, it may be difficult to appropriately respond with a simple "yes" or "no." The instructions further state that the investigator may enter "unknown" if unsure of the answer.

Figure 5. A final fatigue related screen asks if the investigator believes fatigue may have been a factor in the accident.

To further probe what an investigator does during a mishap investigation, we conducted a task analysis on three AF safety personnel who investigate mishaps. The next section describes our methods and results.

METHODS

A more detailed description of the usability testing methods used here is available in the companion technical report by Miller, Eddy, Moise (2008).

REQUIREMENTS ANALYSIS

To understand the role of an investigator who might be responsible for determining the effects of fatigue in the accident investigation process, we used a task analysis (TA) approach, fashioned after Greenberg (2004). Our subject matter experts (SMEs) were three active duty USAF officers (two Majors and a Captain) and the chief of ground safety for an Air Force base. All had served on Air Force Safety Investigation Boards (SIBs) and two of them had taught courses in accident investigation. The three officers were familiar with fatigue management principles.

The requirements analysis session with the three officer SMEs was conducted on 8-9 May 2006 at Brooks City-Base, in a classroom at a conference type table. At one end of the table, a screen displayed information as it was recorded into a digital document.

The SMEs gave brief overviews of their roles in accident investigation, how they became involved in an investigation, and who might be the primary users of our new interface. They then provided details of the knowledge and skill levels of likely users. They listed their goals and processes for investigating a mishap along with the tasks that accomplished them. At the end of the day, the data presented on an overhead screen and recorded throughout process were reviewed and enhanced by the participants. Each step was also reviewed for its possible benefit from some form of computer automation. Where could an automated fatigue assessment tool fit into the HF portion of Accident Investigation? Were there places where supplementary information would be helpful to the investigator? Where should the tool's required inputs be entered? What should the tools outputs be? Should the tool support "what if" capabilities?

On the following day the ground safety chief described ground mishaps and their investigation. He indicated that the same general goals, processes and tasks were used in ground mishaps, but that the depth of investigation was not as great.

WALK-THROUGH

From the user task descriptions and the requirement assessment (RA), we created scenarios for user testing and review. In the walks through the draft interface, the designers, potential end users, and evaluators worked together to step through typical tasks for which the interface was designed. Because the users were familiar with fatigue assessment in mishap investigation, it was not necessary to train them in the concepts of fatigue management. Instead, the group discussed some of the differences between how a model of fatigue differed from a set of rules used to evaluate the presence of fatigue in a mishap. The designers presented the capabilities of the browser-based tool called Fatigue-Performance Assessment System (F-PAS) and its draft mishap investigator interface. The designers demonstrated the tool by entering data for a fictitious scenario. Questions and discussions accompanied each screen of the tool.

We estimated roughly the total number of usability problems in the interface from the number of problems (E) identified during the walk-through. Assuming a detection rate of about 30% in the walk-through, the total number of problems would be about equal to E divided by 0.30 (Bailey, 1997).

INSPECTION EVALUATION

In the inspection evaluations, representative end users tried to do typical tasks with the interface while a designer/evaluator watched, listened, and took notes. We wished to identify usability problems, collect quantitative data on participants' performance and determine the participants' satisfaction with the product. More specifically, we wished to learn:

- Could the test participants complete the relevant tasks successfully?
- How long did it take the participants to do each subtask?
- Did the participants perform well enough to meet their usability objectives?
- How satisfied were the participants with the interface?
- What changes were needed to make sure that the interface would enable more users to perform more successfully?

These latter questions were addressed through the questionnaire shown in Appendix A. The observer collected detailed data using the scoring sheet shown in Appendix B.

We kept in mind the finding that performance and preference do not always match. For about 2/3 of users, performance and preference measures agree with each other. That is, they either perform well and like the Web site or perform poorly and dislike the site. However, performance and preference measures do not agree for about 1/3 of users. They either perform well and dislike the site or perform poorly and like the site. "Various reasons have been proposed for why people often rate a site more highly than their performance would lead us to expect. They may blame themselves for the problems that they have, rather than the site. They may think they would be hurting our feelings by giving the site a low score. They may not be aware of the problems they had and think they were successful when they were not." (Usability.gov).

In this inspection evaluation, we combined four approaches: walk-through and guideline-, heuristic- and scenario-based reviews. By combining the four approaches, we hoped to achieve at least a 50% usability problem detection rate. Assuming this rate, six evaluators were needed to detect 98% of the problems $[1 - (1 - 0.50)^6 \sim 0.98$; Bailey, 1997]. This group included three SMEs and three designers and evaluators.

Specific Scenario and Checklists

The forms used in the inspection evaluations are shown in appendices. Appendix C shows the scenario that the SMEs entered into the interface. It is referred to as the Ramstein scenario. It follows a crew on an airlift mission that originates at Travis AFB, CA, transits Randolph AFB, TX, and enters crew rest at Pope AFB, NC. In the next duty cycle, the crew flies from Pope AFB to Lajes AB in the Azores and enters crew rest. In their final

crew duty period, the crew flies from Lajes AB to Ramstein AB, Germany, where a mishap occurs.

Appendix B shows the data collection form used by the observers during the inspection evaluation. The form was designed to collect objective data about the elapsed time used by the SMEs in their walk-through of the scenario and:

- Number of subtask assists: When the SME cannot proceed on a subtask, the observer gives direct procedural help to allow the test to proceed.
- Number of subtask errors: Instances where the SME had to attempt portions of the task more than once.
- Number of subtask reversals: Number of times the SME had to "back up" to find something on a previous screen that they needed on the current screen.
- Subtask completion (Y/N): Yes = complete and correct achievement of subtask goal.
- Problem severity (0/1/2): 0 = no problem; 1 =minor (users are annoyed, but this does not keep them from completing the scenario); 2 = show stopper (if we don't fix this, users will not be able to complete the scenario; and/or many users will be frustrated if we don't fix this; they may give up).

Appendix A shows the questionnaire completed by the SMEs at the end of their work with the interface. It was used to:

- Rate the ease of application of the Mishap Investigation Tool (MIT) to the intended task: the simplicity with which the MIT could be employed to determine whether fatigue was a factor in a mishap.
- Rate the performance of the MIT: the speed with which the interface responds to requests.
- Rate the support information for the MIT: the information available to acquire, use and support the MIT.
- Rate the MIT's function: the overall capabilities of the MIT.

Additionally, the following open-ended questions were asked of the SMEs:

- What were your objectives as you tested this interface?
- Was the scope of the usability testing that you did adequate to meet your objectives?
- Can you suggest another method of raw data entry that would reduce time, prevent entry errors, and provide greater awareness of data conflicts/errors?
- Can you suggest other data editing methods that would work on a web page and would be more powerful for making changes?
- Can you suggest other ways of displaying the mission log that would facilitate understanding, clarity, and accuracy of the situation?
- Could the MIT report be formatted differently to better assist you in completing your mishap investigation and report?
- Could the MIT graph be formatted differently to better assist you in completing your mishap investigation and report?
- What other improvements should be made to the MIT?

Test Sessions

The first inspection evaluation was conducted at Brooks City Base, Texas, on the mornings of September 6 and 7, 2007. The SMEs were three Aerospace Physiologists who were trained and experienced in mishap investigation and thus familiar with the AF Safety Center's AFSAS database. The MIT interface was modeled in part upon the AFSAS interface.

The observers arrived at the testing site at 0745 on 6 September 2007 and attempted to set up three computers for a scenario walkthrough of the new web-based MIT. Unfortunately, the web-based software was developed under Internet Explorer (IE) Version 7 and the AF was using Version 6. This lack of backward compatibility in the software development tool prevented the entry of new data and the manipulation of old data, specifically data dealing with locations and events. We then spent about two hours showing the tool to the SMEs, but no usability data were collected: the SMEs were walked through a previously entered scenario to familiarize them with the MIT.

Subsequently, we pursued two solutions to the version problem and then met again the following morning to conduct the usability test. Our USAF host had investigated the acquisition of computers with IE Version 7 and we had made modifications to the application that would allow it to operate under IE Version 6. On 7 September, computers with Version 7 were available for the usability test. Only the Save and Print functions of the application were unavailable for the test.

The second walk-through and inspection evaluation was conducted using the Ramstein mishap scenario. It was conducted in the computer laboratory of the altitude training group of the USAF School of Aerospace Medicine (USAFSAM/ATA), Brooks City-Base, Texas, on the morning of 11 September 2008. The SMEs were three flight surgeons from the USAFSAM Residency in Aerospace Medicine (RAM) program, a program in which mishap investigation is taught. They walked through the scenario in parallel, with one observer taking notes.

RESULTS

REQUIREMENTS ANALYSIS

The SMEs indicated that the investigators who might use the tool would be flight surgeons, life support technicians, human factors consultants, and the maintenance member of the board. The maintenance member reports on the personnel and the training of the personnel who maintained the aircraft prior to the mishap. They might also use the tool to investigate a person who may have manufactured, inspected, rebuilt, or installed a defective part or performed a service check on some part or system that failed and caused the accident. In fact the tool might possibly need to include the capability to analyze the shiftwork schedule a person was working at the time of interacting with the defective part. Other primary users might be the Board President, Pilot, other board members or consultants like an aviation psychologist. The SMEs felt the Judge Advocate General officer and maintenance member should be excluded from the interface design requirements.

In discussing the skills and abilities of our users the SMEs felt that users would be E-5s or higher. Investigators completing USAFSAM's AMIP course are trained through review of 17 case studies. The board president attends a specific board president's course and the investigating officer, pilot member, and maintenance member attend different safety courses. In these courses they learn to interview, put their technical training to use, to observe, and to diagnose. They use checklists to avoid failing to ask all of the necessary questions. They indicated that we should make the interface as simple as possible because some users will have no human factors training or fatigue management training. However, they will all be computer literate. Security will be important because of privacy and privileged information issues. They suggested that we include content that has links and phone numbers for help (consultants, possibly at the Safety Center). They suggested that emulating the AFSAS interface would be ideal. They indicated that their human-factors input to the SIB report is mostly in narrative form since the information is not requested by the AFSAS software.

Generally our SMEs had been added to the SIB within 1-10 days after it was formed. They indicated that every mishap is different and that different amounts and types of information has been recorded when they are called to participate.

The SMEs indicated that as they started to become involved in the investigation process, some investigation had already been conducted. Generally, medical records had been collected and an interim report had been generated by onsite people. Although the medical records should report pre-existing sleep pathology, under-reporting is a significant issue in the aviation community because of fear of disqualification from aviation duties.

The results of the requirements analysis, including the automated components, are shown in Table 1. Each of 10 goals and processes was provided by the SMEs. The information was displayed on the screen for all to see as it was discussed and enhanced.

Table 1. Task Analysis with Automated Components.

	PROCESSES/GOALS STEPS	TASKS TO ACCOMPLISH	SUGGESTED AUTOMATED COMPONENTS
1.	Acquire Big Picture/Error	Review ISB and other available data (72 hr history)	
2.	Determine "person/s"	Follow link from AFSAS	
3.	General assessment of medical review	Look for general & sleep pathology, toxicology, medication & hygiene.	List of sedating meds & meds interfering with sleep. List of good/bad sleep hygiene practices.
4.	Assess mishap operator personality, behaviors, skills, currency, work/sleep schedule/patterns including off duty.	Interview witnesses & participants	If fatigue suspected as precondition, provide cue for collecting sleep/work data.
5.	Assess Ops tempo	Review schedules, interview. Sortie generation rate.	Need workload data & model.
6.	Review ORM program & organizational climate	Review specific MSN and overall program	Link to fatigue ORM
7.	Review MSN/task (work cards) profile	Compare with other similar MSNs. Interview other crew with similar MSNs. Assess flight conditions at mishap	Preliminary fatigue analysis. Work/sleep crosscheck. "What if" comparisons.
8.	Assess Environmental Conditions, Sleep	Visit quarters/facilities	Modify preliminary fatigue analysis.
9.	Acquire 1st hand experience of MSN or task	Fly MSN, visit site, simulate. Observe task being done.	
10	Assess operator/maintainer performance	Reconstruct MSN thru simulation. Interview other crew concerning performance of individual. Review recorded data (HUD, CVR, cameras). Look for fatigue symptoms.	List of fatigue symptoms.
Notes: This Table starts with the first steps in accident investigation and progresses down through to the end. Data in each column move from left to right, from global Process/Goals, to specific Tasks that are directly related to them. The last column addresses how an automated tool, such as FAST™ might facilitate the tasks.			

As an investigator accomplishes their goals by completing the tasks listed in Column 2 of Table 1, they need to access information that may not be available in AFSAS or other AF documents or checklists. Some of these items were listed by the SMEs in Column 3, associated with a particular task. Much of this information can be provided by our tool through the mishap investigator's interface. For example, when investigating an individual's medical records it would be helpful to compare the medications they are taking with a list of sedating medications and medications that might interfere with sleep.

If any of these medications were used, the time of administration could be important to determining the cause of the mishap. Other items listed as important supplementary information included a list of good/bad sleep hygiene practices and a list of fatigue symptoms that might cue interviewees to provide information of that kind.

The general consensus of the SMEs was that, whatever form the tool took, it should be easy to use by everyone. They suggested that since users would have had familiarity with AFSAS and its data entry screens, the new tool should follow that format as much as possible. This would also eliminate the need to learn two separate programs. In an ideal world, the new tool would interface with or be embedded within AFSAS to avoid the need for double data entry. The following summarizes the findings regarding a fatigue decision-aid tool. It should:

1. Follow the AFSAS format for data entry and editing whenever possible
2. Have hot buttons or pull down menus for accessing supplementary information such as a list of sedating medications, stimulating medications, good/poor sleep hygiene practices, and fatigue symptoms
3. Provide a list of questions for obtaining sleep times from interviewees and spaces for recording the information compatible with the data entry format
4. Provide fatigue analysis in an ORM format as well as using current fatigue indicators
5. Provide a preliminary fatigue analysis
6. Provide a work/sleep crosscheck to alert the user to overlapping work/sleep times
7. Provide "What if" comparisons so that the current mission can be compared to other similar missions. Also, compare with changes to the mishap mission schedule, such as delaying takeoff 2 hours.
8. Maintain multiple versions of analyses that can be labeled differentially

INTERFACE PLANNING

Based on our research of AF documents, we found no forms for collecting mishap data. Since the recording of accurate sleep times, Item 4 in Table 1, is necessary to obtain a reliable fatigue analysis, we recommended a format for recording the information. Our recommendation was to use a time-based, "sleep/activity log" when conducting investigative interviews. Using these logs, the investigator would be able to record a much more detailed picture of the actual fatigue state of any individual involved at the time of the incident. Using a simple paper-based log like that in Figure 6, it would be possible to precisely record sleep times, activity times, and subjective measures of sleep quality for as much time preceding the incident as data were available. In the following example of such a log, actual clock times corresponded to the start and end of sleep or work, while the type of activity engaged in by the individual at the time was recorded in the bottom row (O = "on duty", T = "trying to sleep", S = "sleeping"). Sleep quality could be scored as Excellent, Good, Fair, or Poor. Such a log could be a modifiable blank electronic template (with ability to print hardcopy versions) and included as part of the fatigue analysis tool.

Mishap <u>123456_____</u>
Investigator <u>Isaac Investigtor</u>
Crew Member <u>Pilot_____</u>

Zulu	12	13	14	15	16	17	18	19	20	21	22	23
Local time	7	8	9	10	11	12	13	14	15	16	17	18
Actual Time	1212						1855			2105		
Sleep Quality										fair	good	good
Activity	O --	-- --	-- --	-- --	-- --	-- --	-- O			T S	-- --	-- --

Figure 6. This activity log may be used by investigators for recording data from interviewees.

This type of log may be quite useful as the investigator might collect numerous accounts regarding when the individual in question may have slept during the days leading up to the incident. It is our opinion, and that of the SMEs, that this type of information should be gathered on <u>all</u> mishaps for evaluation to cue the investigator as to whether fatigue might be a precondition. Since fatigue can contribute to any type of unsafe act (e.g., skill-based error, judgment error, perceptual error, violation), there is nothing specific about an incident which allows an investigator to include/exclude fatigue as a potential precondition. The modeling tool itself should be used as a fatigue-screening test. The sensitivity of an investigator to fatigue should not determine whether confirmatory testing with F-PAS is necessary; it should be done routinely. In the opinion of one of the SMEs, that is why fatigue is presently so under-diagnosed in mishaps. Another function these logs allow is that they may be corroborated with one another, checked against known duty rosters, and used to establish an accurate timeline of the individual's sleep hygiene leading up to the incident.

These data could be summarized for data entry using the format of Table 2. Using the AFSAS data entry screens as a model, we could create similar data entry forms for the F-PAS tool. Note that multiple sleeps within a day can be accommodated in such a summary table. Ultimately, this log and summary table would allow the investigator to enter accurate information into the corresponding AFSAS or F-PAS data input interface.

Table 2. Summary Table

Person/ Date	Times in Zulu				Daylight Savings	Quality
Pilot	Sleep start	Sleep end	Duty start	Duty end		
23 Oct 08	2105	0430	1212	1855	Y	Excellent
24 Oct 08	2125	0425	1200	1900	Y	Good
-						

The above describes how data could be collected and summarized. The next question was how to begin a data entry session using the F-PAS tool. We needed an opening screen that would document general information concerning the mishap and tie the final fatigue analysis back to the AFSAS mishap report. As an opening screen, we settled on using

something like the AFSAS data entry screen "Create Report," Step 1 shown in Figure 7.
In addition to the information requested by AFSAS in Figure 7, our screen would request
two additional pieces of information. First, it would ask for the Mishap Number so the
fatigue analysis could be easily associated with the AFSAS report. Second, it would ask
for the date and time of the mishap in the time zone of the person rather than the vehicle.
The pilot of a UAV may be many time zones removed from a mishap involving the UAV.
Third, the model would need the latitude and longitude (lat/lon) of the person involved in
the mishap. The lat/lon would be requested in the format shown in Figure 8 for
compatibility with AFSAS, except that seconds could be omitted from the specification of
lat/lon. This information could be used to create a file name for the schedule using the
individuals name and number. The mishap would be entered as a special schedule event.

Figure 7. Fatigue analysis start up screen.

**Figure 8. AFSAS screen showing the format for the latitude
and longitude of the mishap.**

Figure 9 shows a schematic of an Entry Screen for starting the data entry. Time of Mishap would be entered using the 24-hour clock. Mishap Location would be the location of the person involved in the mishap. Days Prior to Mishap would default to two weeks so that a history of sleep, work, and important events for the person could be entered prior to the mishap. If less than two weeks of data were available, a number of days less than 14 could be entered.

Fatigue Analysis
Step 1
Initial Entry Screen

Mishap #			
Name			
Time base	Zulu/Local		
Time of Mishap			
Mishap Location			
Latitude	Deg	Min	N/S
Longitude	Deg	Min	E/W
Daylight Savings Time	Y/N		
Days Prior to Mishap			

X   Enter Sleep Data

Figure 9. Entering mishap for fatigue analysis would start with basic information shown in this schematic of the start screen.

Once the basic information about the mishap and its date was known by the system, information on work and sleep for the days preceding the mishap could be entered. Although our model would need to ask more in-depth questions than are asked currently by AFSAS, these questions could be accommodated in the format of an AFSAS data entry screen similar to Figure 4.

An entry screen for sleep history could contain the elements shown in Figure 10. An option to consider would be for various fields that are likely to be the same from sleep period to sleep period to carry over from the previous screen. The repeating fields would be Time Base and Daylight Savings. The date field would decrease by one. If all the defaulted fields were correct, the investigator would merely add the start times of sleep, end times of sleep, and select the sleep quality.

Fatigue Analysis
Step 2
Sleep Entry Screen

Time Base　　　[X] Zulu　　　[ ] Local

Mishap NUMBER, DATE and TIME

Enter each Sleep Period Working back in time from the Mishap

Start		End		Daylight Savings	Quality
Date DDMMMYYYY	Time HHMM	Date DDMMMYYYY	Time HHMM	Y/N	Excellent, Good, Fair, Poor

If sleep is interrupted by more than 15 minutes a new sleep period should be entered.
Sleep periods are not needed for more than two weeks prior to the mishap.

[X] Continue Entering Sleep　　　[ ] Review Sleep History

Figure 10. This figure shows a schematic representation for entering sleep history.

Once the Review Sleep History box at the bottom of Figure 10 was checked, the sleep history summary would be shown similar to Figure 11. It would contain the information in the format of a spreadsheet.

Fatigue Analysis
Step 3
Sleep Summary Screen

Time Base　　　[X] Zulu　　　[ ] Local

Mishap NUMBER, DATE and TIME

Review each Sleep Period and Correct as Necessary

Start		End		Daylight Savings	Quality	Delete Period
Date DDMMMYYYY	Time HHMM	Date DDMMMYYYY	Time HHMM	Y/N	Excellent, Good, Fair, Poor	
						[ ]
						[ ]
						[ ]

[X] Enter New Sleep Period　　　[ ] Enter Duty Periods

Figure 11. This figure shows a chart for summarizing the sleep history.

The sleep history table could be edited if needed or the investigator could enter the sleep information directly in this format rather than one sleep period at a time. If an additional sleep period was needed, the investigator would check the Enter New Sleep Period box to enter the new information. Once all the new or additional sleep periods were entered (See Figure 10) and the Review Sleep History box was again checked, the sleep history table would be displayed for review. Sleep periods could be completely deleted using the check box to the right of each row.

Once the Enter Duty Periods box was checked at the bottom of the sleep history table (Figure 11), the investigator could begin entering the duty periods for the individual similar to sleep. When complete, the duty history would be reviewed and the investigator could move on to events. However, in our attempts to assemble the interface display screens to meet the SMEs requirements, we came to a problem that was created by our approach of using the AFSAS format. We became concerned that the process of reviewing and verifying the data that had been entered might lead to data entry errors. Tables of times, dates, and lat/lons are just strings of numbers and numbers can be reversed, misread, and easily skipped. Further, keying strings of numbers would be very time consuming. Several members of the team were gathered together to discuss the problems and possible solutions.

We began by considering a two pronged approach consisting of following the AFSAS format or a Scheduling Grid (SG) approach similar to FAST™. The AFSAS format approach consisted of text entry for each work, awake (occupied), and sleep interval. This would then be followed by a verification phase where the entered data would be compared with the original data collected from the interviewees. The SG approach allowed the investigator to enter the above listed intervals using a point, click, drag, and drop method similar to FAST™. We decided after much discussion that the AFSAS format would be too cumbersome, slow, and error prone to pursue. The problems were:
1. Keying in all the values for up to 14 work periods would take considerable time.
2. Keying in all the values for occupied periods when the subject was neither on duty nor sleeping would take considerable time.
3. Keying in all the values for time zone crossings would be separate from the work intervals that might lead to loss of the big picture.
4. Keying in all the values for sleep periods would take considerably more time. Because of the sheer volume of awakenings (interruptions during sleep), many would likely be left out reducing the accuracy of the tools projections.
5. Summaries of the above information would be difficult to cross check against the original data (verification). Comparing lists of numbers for accuracy is a difficult and error prone task for a human.
6. Without some graphical presentation, getting the "big picture" of work and sleep leading up to the mishap would be difficult to conceptualize.
7. Overlap of sleep and work intervals and coordination of time zone changes with time of occurrence would be difficult to recognize from tabular data.

After consideration of several options, we decided on a SG approach that would allow the investigator to enter nearly all the information using a point, click, drag, and drop method similar to FAST™. We presented the two approaches to the investigators along with the pros and cons and let them decided on how to proceed. The investigators agreed that the SG approach would be faster and provide a format that would be easier to verify.

For all of the above reasons the SG approach was pursued even though some training of users might be required. The SG approach was considered more compatible with a time scale (a timeline) format that lends itself to what happened before the mishap that may have been causal. Further, transferring data from a paper data collection form to an entry screen would be very easy and the graphic layout would facilitate error recognition.

The approach adopted closely follows the Activity Log used by sleep researchers. After the investigator completed the preliminary schedule data entry (date, location, etc.), a scheduling grid would be presented for the sleep and work data. The size of the grid would be based on the number of days of data available to the investigator. It would contain blocks of four rows beneath Zulu time in 1-hour columns. Each block would represent information for one 24-hour period. The four rows would be local time, event, latitude/longitude (lat/lon), and activity, similar to a paper form used to collect the data from interviewees. The paper forms would contain data in two, 12-hour periods with additional data collection areas below the rows that would hold detailed information concerning the events (type, description, notes, lat/lon (if needed), etc.). Event types would be Show Time, Takeoff, Arial Refueling, Mission Event, Landing, Post-Brief, Mishap, and Misc. It would be necessary to use two, 12-hour screen entry blocks for compatibility with the paper data collection form.

Events and lat/lon coordinates would be entered through windows similar to FAST™ for entering such information (text and pull down menus). Local time would be recomputed after events containing a lat/lon were entered. The local time at the lat/lon would be displayed in a pop up window so the investigator could crosscheck it with his or her knowledge of the local time at that location. The computer-generated local time could differ from that determined by the investigator for two reasons. The data entered for lat/lon could be in error or the investigator may have made an error in determining the local time at the lat/lon. In either case, the conflict would need resolution before the data could be analyzed.

Next, work, occupied, and sleep intervals would be entered using the mouse. A cell representing 15 minutes in a specific hour in the activity row would be selected, a left button press would be initiated, the pointer would be dragged to the end of the interval, and dropped. A window would open to select work, occupied, or sleep. If sleep were selected, another window would open asking for the quality of sleep. A small window would show the Zulu and local time under the pointer during the drag operation. The investigator would enter all the work, occupied, and sleep/nap intervals for the 24-hour period.

Each 24-hour period would be entered on a separate screen for each date. A grid similar to FAST™ would be shown for verification after all data were entered for all days (minimum

3 days). Moving the pointer over an event would show the details of the event. Moving the pointer over an interval would show the details of the interval. Events and intervals could be edited from this review screen. Once the data were verified, they could be submitted for analysis.

WALK-THROUGH

A meeting was held with the three officer SMEs who had been involved in the RA on 22 June 2007 at the same location as above. The purposes of the meeting were to review progress on the Mishap Investigation Tool and to receive recommendations from the SMEs. The initial draft of the interface was demonstrated by entering data for a fictitious scenario, then making changes to the data and observing the results. Questions and discussions accompanied each screen of the tool. The following points came out of the meeting:

1. Add text to the mishap number entry question to the effect that this should be the AFSAS mishap investigation number issued by the AF Safety Center.
2. Check for a valid mishap number by verifying that the length does not exceed 6 digits.
3. Get the official list of mishap roles from the Safety Center.
4. Use an intelligent prompter (like Google) when looking up cities or bases. This was a good suggestion that presumably will make the search time through the city/bases lists quicker.
5. Use the nearest base as location input and allow entry of more exact latitude/longitude values. Test for valid values by allowing a radius around the city coordinates.
6. Enter Dates as YYYY MM DD.
7. When time entry is not exact indicate such with a phrase like "What is the nearest Zulu time?"
8. Correct the typical sleep time question by adding "for the individual."
9. Add the phrase "Not Recommended" to the "change typical sleep time" question.
10. Indicate that coordinates in the Event window can be edited.
11. Check for mismatched location data entered.
12. Provide generic feedback when entering data: A) changing locations B) events during sleep C) sleep during a work period.
13. Rephrase Lapse Index (LI) on report: add 1.0 to make the value >= 1.0
14. How were the cutoff values for the report warning flags determined?

INSPECTION EVALUATION

First Inspection Evaluation

The specific usability comments of the SMEs were coded and organized into Table 3. The full list is available in Appendix D. The items come from the first 32 steps and the 39th step in the data sheet (Appendix B). Due to the time spent in analysis of the preceding steps, steps 33-56 (entry of second crewmember sleep data) were not performed. The comments are separated into what was being inspected the interface, the worksheet, or the help file. The primary problems involved terminology and confusion among choices. The

SMEs had excellent recommendations for improving the content presented on the screens and on the worksheet.

Table 3. Table of Investigator Comments.

Comment Category	Interface	Worksheet	Help File
Confusion:			
selection among items	4		
what next	2		
terminology	5	2	1
User error: missed element		2	
Software error	2		
Recommendations	10	8	1
Positive comments	1		

The elapsed times spent by the SMEs entering data on the Mission Log worksheet were 23, 36 and 35 minutes (mean = 31.3 minutes). The elapsed times spent by two of the SMEs entering data in the interface were 24 and 27 minutes.

The SMEs had some difficulty converting local time to Zulu, since they did not do it often enough to stay practiced. It was decided that prompts on how to find a Zulu calculator on the web should be made available to the investigators. If we gave them a button or URL in F-PAS, it would always have to be updated to make sure the site was always there. Also, there will be a general, tabular guide to time zones around the world printed on the reverse side of the Mission Log worksheet.

Because of the large number of software changes needed and the general agreement that a second usability test would be required, the use of the ratings and discussion questionnaire (Appendix C) was postponed until the next test.

Planned Software Changes

We decided to change the entry of sleep and work intervals similar to the entry method for events. Clicking on the Activity line would open a panel for data entry and would give a default date and start time associated with the hour and 15-minute block where the cursor was pointing at the click. They could enter a sleep or work interval by keying in any two of the following: start time, duration (hours and minutes), end time. The third would be computed by the software and automatically entered into the remaining field. After the changes were made, the user could click:

- Ok, to enter the data into the activity line and close the panel,
- Clear, to delete the default time, or
- Cancel, to close the panel without making any changes.

Once a work or sleep interval had been entered into the data entry screen for the day, clicking on the activity line would have different effects depending on where the user clicked.

- Clicking on unscheduled time would allow data entry similar to what is described above for work or sleep intervals.
  - The data entry panel can never cover the activity line so that the user can see the previous entries while entering a new interval.
- Clicking on a sleep or work interval would allow either (1) editing of the interval or (2) insertion of an interval within the existing interval (entry of sleep within a work interval, entry of work within a sleep interval, or clearing an interval within a sleep or work interval). The user would select one option or the other from a menu.
- Editing an interval would display the same panel as what is described above for work or sleep intervals, but with the start, duration and, end fields filled in. The user could change one or more of the three fields.
  - Changing the start time would change the end time to keep the duration the same (equivalent to dragging the interval)
  - Changing the duration would change the end time to be consistent with the old start time and the new duration (equivalent to lengthening or shortening the interval from the start time)
  - Changing the end time would change the duration, but keep the start time the same (equivalent to lengthening or shortening the interval from the start time)
  - To change the start time and duration, but keep the end time the same, the user would first change the start time and then either change the end time back to the original value (shortening or lengthening the interval while maintaining the end time is assumed to be a less likely operation)
- Inserting an interval into another interval would display an enhanced panel showing the existing interval and 3 fields for data entry of the new interval. The start time of the new interval would show a default time associated with the hour and 15-minute block where the cursor was pointing at the click. The user could enter a sleep or work interval by keying in any two of the following: start time, duration (hours and minutes), or end time. The third would be computed by the software and automatically entered into the remaining field for the new interval. The inserted interval could be surrounded by the old interval in which case the old interval would become two.
  1. The inserted interval could extend beyond the end time of the old interval in which case the old interval would be shortened to the start time of the new interval.
  2. The inserted interval could extend before the start time of the old interval in which case the old interval would be shortened to start at the end time of the new interval.
  3. The inserted interval could completely engulf the old interval in which case the old interval would cease to exist.
  4. No other options would be allowed.
  5. If the user entered the same type interval within the old interval, the old interval would be lengthened if #1 or #2, or the same if #3.

Other planned changes were:
- On the introduction screen, say "Enter Data and Do Analysis" instead of "Do Analysis."

- Instead of "New Analysis," say "Start Over" to show that all data will be lost.
- Alphabetize the mishap participant list.
- Among the Questions pages, the role question is skipped on its first instance. In subsequent executions of the Questions pages the questions are presented correctly. This problem seems to be related to operation of F-PAS on the server.
- During server file browsing a file operations error occurs the first time the server directory is browsed. Subsequent browsing works correctly. This problem seems to be related to operation of F-PAS on the server.
- On the Mission Log pages, activity data entered on days -2 and -1 are sometimes lost.
- On the Mission Log pages, the software crashes when entering data which overlaps days (specifically days -3 to -2) and then moving to the subsequent day (-2 in this case).
- The start of default sleep should be 2200 and not 2300.
- The number of event types is too small.
- The time in the data entered list needs to have "L" or "Z" appended.
- If user clears the mishap event, F-PAS should reinstate it if the user reenters the questions page.
- The work and sleep interval data entry screens need to have "L" or "Z" appended to the times.
- The work and sleep interval data entry screens need to state that the user is to enter any 2 of the following: start time, end time, duration.
- In the Report, change text to "The Mishap Investigation Tool evaluated __ days of work and sleep data."
- In the Report, change the report's width to better cut and paste into Word documents.
- Mishap time needs an "L" or "Z" at the end.
- Include latest drug data and algorithms.
- Include transmeridian AutoSleep algorithms.
- Add a graphing capability to MIT.
- Allow the optional selection of the day (relative to the mishap day) in which data will be entered (day -6 for example).
- Add on-screen instructions to the Mission Log page.

Summary of Beta Interface

In consideration of concerns expressed above, the mishap investigator's interface was completed and has the following flow. After signing into the F-PAS website with user identification and a password, the user may select from the mishap-investigator interface options to download and print a data collection worksheet called the Mission Log Worksheet or create, edit, or analyze a schedule. The worksheet is a PDF file and can be printed and copied for collecting sleep, work, and event data. Its format matches that of the data entry interface making data entry from the sheet easy, straightforward, and hopefully error free. If a new or previous schedule is selected, the user may enter or edit schedule data. These data fields describe the schedule and provide an association to the AFSAS report. For example, there are fields for who is under investigation (e.g., pilot,

copilot), who the investigator is, the time, date, and location of the mishap, etc. Dates and times may be in either a single local time or Zulu easing the burden of conversion. Once these data are entered, the investigator begins entering the work and sleep schedule of the mishap participant. This information along with events such take offs, landings, etc. are entered into a daily timeline. In this way, the relations among these data elements can be clearly observed so that errors are obvious and easily corrected. The investigator works backward in time from the mishap one day at a time. The software allows entries that span days. All AF bases and many cities are available in a database that contains their lat/lon coordinates saving the investigator lookup time. However, lat/lon entry is also an option for out-of-the-way places that are not near a base or city. The investigator may enter as many days as s/he has data, but three days is a minimum for an accurate fatigue analysis. The investigator may save the file at any time and return later to add more data or make corrections. The investigator may save data files on the server and then copy them to the investigator's computer using the File Management functions available from the F-PAS Main Menu.

Once the data are entered, the investigator selects a button to analyze the schedule of sleep. The result is an analysis of fatigue at the time of the mishap. The performance effectiveness of the mishap participant is presented along with the status of five fatigue indicators. Also shown are the values of the fatigue indicators and their fatigue criteria. For example, if wakefulness in the last 24 hours exceeded 17 hours, the status flag for that indicator shows red. In addition to the fatigue indicators, other common variables computed with the model are shown such as the lapse index and predicted reaction time. An investigator may request a table showing performance effectiveness for each work interval, while awake, at all events or see schedule statistics such as the number of work intervals, etc. Two other options for the investigator include a graph of the entire schedule showing performance effectiveness relative to dates, times, and events, and a report describing the schedule and its fatigue impact. Both may be in either Zulu or the selected local time base. All displays may be printed or saved to a file. An investigator may return to the mission log to correct data or to the preliminary questions. The investigator may also save the file under different names and modify the schedule so that "what-if" type questions can be asked about various schedule variations that might have avoided or reduced fatigue. This is an important option since a mishap investigator is trying to discover not only what may have contributed to the mishap, but also how it might have been prevented.

Second Inspection Evaluation

The three flight surgeon SMEs were Majors with 7, 8, and 14 years of active duty. One had used FAST™ for mishap analysis, one had been exposed to FAST™, and one was completely unfamiliar with FAST™. They walked through the scenario in parallel, with one observer taking notes. The SMEs' comments about the MIT follow.

Mishap Questions page:
- In the pop-up window that allowed the user to specify the location (city, base or lat/lon) of the mishap, a user did not understand the abbreviation "DST." Recommendation: state the meaning, "Daylight Savings Time," explicitly or

nearby or by hyperlink.
- The question about the location of the mishap operator was too vague. It should be more specific. Recommendation (example; italics added to show difference): "Where was the mishap copilot *physically located* at the time of the mishap?"

Mission Log pages:
- The users could not locate the "hot" areas on the Mission Log table (i.e., the event and activity rows) to click with the cursor, even though they read the relevant instruction in the middle of the page. Recommendation: Change the cursor shape over the hot rows.
- The difference between the row labels "Event" and "Activity" was unclear to the users. One SME noted later that the graph used the designation "Sleep/Work" instead of "Activity." Recommendation: Use the designation "Sleep/Work" instead of "Activity" for the bottom row of the Mission Log table.
- In the pop-window that allowed the user to specify an event, a user did not understand abbreviation "DST." Recommendation: state the meaning, "Daylight Savings Time," explicitly or nearby or by hyperlink.
- Because the start of a sleep period or work period was not a defined event, the users perceived that they were unable to specify the location of the mishap operator at the beginning of the F-PAS schedule. They failed to use the "Other" option for such an event. Possible recommendations: 1) add beginnings and ends of sleep and work periods as explicit events, 2) create a new question in the Questions section where the user can specify the location of the earliest point in the schedule
- The lack of feedback on the Mission Log display about event times caused the users to doubt that their data entry for events had been accepted properly by the software. Recommendation: When entering data in Zulu, show the actual, local time of the event on the Mission Log's grayed-out row for local time.
- It was unclear to the users that when entering new sleep or work periods that they could enter *either* duration *or* end time. Recommendation: State this explicitly.
- The users noted that, when entering a sequence of events such as takeoffs and landings, the pop-up window always defaulted to the original location used during software start-up and that this was confusing and irritating. Recommendation: Use the immediately preceding location in the schedule's timeline as the default location for a new event.
- Two users were unclear about the meanings of the colors used in the Activity row of the Mission Log table (blue, red and gray). Recommendation: Define these colors' meanings on the Mission Log screen and in the Help file.
- When the users edited sleep intervals, they found that sleep qualities that had been set previously to Good and Fair were reset to Excellent. Recommendation 1: Sustain the original data entry. Recommendation 2: Default the sleep quality choice to a "Choose one" option, and then show a warning if none is chosen. Recommendation 3: Include sleep quality in the hover box associated with sleep periods.
- The users did not understand the meaning of the phrase "Location Included" in the Event data entry window. Checking the box next to this phrase seemed to make no

difference in the display.

Only one SME's data entry reached fruition, i.e., display of the mishap analysis, without software problems on the first try. The other two were halted by "script" errors. The first of these two SMEs was forced to re-start data entry. The other recovered by re-entering some of the data in the Mission Log.

The first SME reported that the mishap event was no longer being displayed on his Mission Log page for the mishap day. The observer directed him in troubleshooting the problem. He displayed Day -1 and then the mishap day again, to no avail. Next, he backed up to the last question and then re-displayed the mishap day in the Mission Log. The mishap event was still missing. Finally, he backed up through the questions to the mishap location page and then moved forward though the questions to the mishap day in the Mission Log. The mishap event was still missing. Without any other recourse, he started over from the main menu.

The second SME had received a script error when trying to enter an event. He clicked "Yes" to end the script and a pop-up window for an event was displayed, but all the fields were blank and the hour was "un." He filled in the fields and the software seemed to accept his inputs; the event was displayed on the Mission Log. Later, when he clicked on the Mishap Analysis button, he received an error message. He backed up with the browser arrows and found a number of errors in the Mission Log, on two different days. These were errors caused by the software, not by the SME's original input. He repaired those errors and was rewarded with a correct Mishap Analysis. However, in the graph he found a takeoff and landing in a sleep period. Upon investigation, he found that the start-up sleep period was missing in the Mission Log. This may have been a software-induced error in the Mission Log that he missed after the Mishap Analysis failed. He re-entered that sleep period in the Mission Log, and the graph appeared correctly.

Mishap Analysis page:
- It was unclear to the users that the analysis referred solely to the time of the mishap. Recommendation: Instead of using the subtitle "Mishap Analysis," use something like "Fatigue Estimates at Time of Mishap."
- One user was concerned about the inability to enter two events within a single hour; for example, a takeoff and mishap that happen only minutes apart. The explanation that a difference of up to 60 minutes would make little difference in the estimates provided by the SAFTE model and, thus, that one could just enter the mishap event in that hour, was fully acceptable to the user. Recommendation: Place this explanation in the Help file.
- The users asked what the reference was for the percentages presented on this page. Recommendation: State explicitly on the page that they refer to a "fully rested individual." Explain more fully in the Help file.
- The users did not understand the need to show the percentage of sleep reservoir. Recommendation: Delete this statistic from the Analysis page.
- One user noted that the Analysis threshold for sleep recency was 17 hours, but that the Report said that sleep should follow 16 hours of wakefulness. Recommendation: Amend the language in the Report appropriately.

29

- The users did not like the functional grouping of the symbols and statistics on the Analysis page. They found it difficult to relate statistics, thresholds, and symbols with one another. They also did not like the empty cells when red flags were not shown. They recommended a four-row display: titles, symbols, statistics, and threshold values. The other statistics could be shown in a vertically aligned table below that (without Reservoir), or just deleted. A mock-up of their suggestion is shown in Figure 12. Additional recommendation: a green flag may imply a safe condition. Instead of green flags for the fatigue indicators, use a white flag, which is neutral.

Fatigue Predictions at Time of Mishap

Variable	Performance Effectiveness	Recent Sleep	Chronic Sleep Debt	Hours Awake	Time of Day	Out of Phase
Prediction	89%	4.5 hrs	7.1 hrs	16.2 hrs	17:35	3.2 hrs
Reference Threshold	> 90%	< 8 hrs	> 8 hrs	> 17 hrs	00-06	> 3 hrs

Figure 12. Mock-up of suggested Dashboard at time of mishap.

Mishap Graph:
- One user was unfamiliar with the green, yellow, red, and dotted line codes in the graph. Recommendation: Explain these codes in the Help file.
- One user requested that he should be able to cut and paste sections of the graph by specifying the beginning and ending with date/time.
- None of the users liked the fact that the line representing the performance effectiveness (PE) value plummeted immediately to zero after the end of the F-PAS schedule data. Recommendation: Leave the line hanging at the PE value at the time of the mishap.
- The users noted that the date/time notations on the x-axis of the graph were "too busy."

Mishap Report:
- The first sentence of the report stated that the next sentence was a summary statement. The users agreed that this was unneeded. Recommendation 1: Use the heading **SUMMARY** before the first sentence, and then use a few other bold headings throughout the report. Recommendation 2: One heading should be for statistics concerning the whole schedule. Another should be for statistics concerning the 72 hours before the mishap.

- The users noted that there was no information about individual differences. Recommendation: Include a statement about the confidence interval around PE.
- The users did not understand the numbers in the "Quality" column of the sleep intervals table. Recommendation: Use text in this column.
- One user perceived a disconnect between an estimated PE of 90% and a 1.4x lapse likelihood. Discussion revealed that he perceived 90% PE as being a much better level of cognitive function than it is in fact. Recommendation 1: Emphasize in the Help file the relationship between 90% PE and 16 hours of continuous wakefulness. Also, link to this explanation from the Analysis page. Recommendation 2: Include a graph of the relationship between PE and lapse likelihood in the Help file.

The SMEs' time constraints precluded testing each individual alone. Thus, no time-lapse data were available. In addition, the SMEs departed the testing session before subjective-ratings data could be collected. All of the problems identified in the second inspection evaluation were scheduled for fixes at the time this report was prepared.

DISCUSSION

REQUIREMENTS ANALYSIS

The mishap investigation process was fully represented to the authors through manuals, documentation, task analysis, discussions, and examples. In addition, one of the authors (JCM) had previously conducted many fatigue analyses for Safety Investigation Boards as a part of the Biobehavioral Performance Branch at Brooks City-Base. While each mishap is unique, the authors were confident that they fully understood how a fatigue analysis tool should fit into the overall investigation process. As the mishap investigation process unfolded, it became clear that there was an additional need for instruction on how to collect accurate sleep data for insertion into the tool. Therefore, the web-based help manual contains a well-documented method for obtaining the sleep times of the mishap participant.

The requirements analysis provided excellent information for designing the MIT. The contributions of the SMEs continually kept the authors on the important issues of the interface. While not all the requested interface features have been implemented at this writing, they will in the coming months.

WALK-THROUGH

Working with a user to enter data from a mishap into the MIT was extremely important for finding interface flow problems and identifying additional information that might be needed. The final walk-through of the MIT exposed inter-screen problems that did not occur in merely discussing the display screens. The walk-through opportunities provided significant and concrete recommendations for specific interface modifications. Designing an interface without several opportunities for the users to walk-through a scenario that exercised the tool's features would miss many important human factor errors. Little details such as date format become obvious in a walk-through.

INSPECTION EVALUATION

Having a user enter data from a mishap into the MIT was extremely important for finding interface problems and illuminating subtle issues. The final inspection evaluation of the MIT exposed software problems that had not occurred in alpha testing. Thinking through the steps in the evaluation process for a specific scenario (mishap) can be enlightening if the designers have not considered how information is brought together for data entry. The inspection evaluation opportunities provided significant and concrete recommendations for specific interface modifications. Because the requirements analysis clearly defined the order of tasks and where they fit into the overall process, there was minimal restructuring of the interface between each inspection. Most of the identified issues were related to labeling, terminology, and better ways to define choices. What appears completely straightforward and understandable to the designers can be confusing or unintelligible to even a knowledgeable user.

The data collected in an inspection evaluation permit the designer to streamline procedures to save the user time and prevent errors. In addition, having the actual times that it takes to enter the necessary information into the tool can be used to convince hesitant, resistant users that the time is worth the effort. Often a potential user will see the benefit of a new

product, but will not want to spend the time to learn how to use the product to achieve its benefits. With data entry times and learning times, the new user can better see the cost benefit of the new product. Output products that are compatible with the users work environment also help to increase usage.

The F-PAS product has additional interfaces for shift work scheduling (Miller, Eddy, Smith, & Moise, 2009) and mission planning (Eddy, Moise, Miller, & Smith, 2009). Its development process has generated other reports that may be of interest to the reader. Miller & Eddy (2009) discuss the use of F-PAS in Operational Risk Management of Fatigue Effects. Eddy & Mendez (2009) report on verifying F-PAS predictions against FAST™.

CONCLUSIONS

We recommend that a fatigue assessment decision aid like F-PAS become a required item in every mishap investigation. Such an objective fatigue assessment tool would standardize the data available for studies conducted in the AFSAS database to investigate overall mishap trends. At this writing, AFSAS only includes sleep data for 72 hours in the database. Further, no metrics are recorded that avail themselves to quantitative analysis. Use of F-PAS and incorporation of the performance effectiveness score and the presence or number of fatigue indicators would greatly benefit future research into the insidious effects of fatigue on human error.

This report has documented the development of a World Wide Web-based fatigue analysis product that was designed for mishap analysis. The effort used task analysis to develop the requirements for the interface and used walk-through and inspection evaluations to assess the success of the endeavor. The walk-through and inspection evaluation processes indicated that most of the requirements of potential users were met reasonably well and that potential users were able to operate the interface reasonably easily. While the overall outcome of the effort successfully produced a useful product, it continues to undergo modifications to improve reliability and to implement all the features identified in the requirements analysis. Had the tool been computer-based instead of web-based, the implementation of the MIT would have been much easier to implement and there would have been more time to complete more of the recommended features.

REFERENCES

AFSAS User's Guide (2006). (AF Instruction No., AFI 91-204). Kirtland AFB, NM: HQ
    AF Safety Center. https://sas.kirtland.af.mil/,
    https://sas.kirtland.af.mil/pls/LOGIN/loginuser.pkg_login.main.

Bailey RW (1997). Usability studies and testing. Chapter 12 in TS Andre and AW
    Schopper, *Human Factors Engineering in System Design*, CSERIAC SOAR 97-03,
    Defense Technical Information Center, Fort Belvoir, MD.

Department of Defense Human Factors Analysis and Classification System (DoD
    HFACS), A mishap investigation and data analysis tool. Kirtland AFB: Air Force
    Safety Center. http://www.safetycenter.navy.mil/hfacs/downloads/hfacs.pdf,
    http://afsafety.af.mil/SEF/Downloads/hfacs.pdf.

Eddy DR & Hursh SR (2001). *Fatigue Avoidance Scheduling Tool (FAST)*. (AFRL
    Technical Report No. AFRL-HE-BR-TR-2001-0140). Brooks AFB, TX: Human
    Effectiveness Directorate Biodynamics and Protection Division, Flight Motion Effects
    Branch.

Eddy DR & Hursh SR (2006a). *Fatigue Avoidance Scheduling Tool (FAST™) Phase II
    SBIR Final Report, Part 1*. (AFRL Technical Report No. AFRL-HE-BR-TR-2006-
    0015). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and
    Protection Division, Fatigue Countermeasures Branch.

Eddy DR & Hursh SR (2006b). *Fatigue Avoidance Scheduling Tool (FAST™) Phase II
    SBIR Final Report, Part 2*. (AFRL Technical Report No. AFRL-HE-BR-TR-2006-
    0040). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and
    Protection Division, Fatigue Countermeasures Branch.

Eddy DR & Mendez J (2009). *24/7 Operational Effectiveness Toolset: Predictions of
    FAST™ and F-PAS Compared*. (AFRL Technical Report, Contract FA8650-06-C-
    6606). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and
    Protection Division, Biobehavioral Performance Branch.

Eddy DR, Moise SL, Miller JC, & Smith R (2009). *24/7 Operational Effectiveness
    Toolset: Mission Scheduler Interface*. (AFRL Technical Report, Contract FA8650-06-
    C-6606). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and
    Protection Division, Biobehavioral Performance Branch.

Gebicke ME (1996). *Briefing Report to the Ranking Minority Member, Subcommittee on
    Military Procurement, Committee on National Security, House of Representatives,
    Military Aircraft Safety, Significant Improvements Since 1975*. (GAO Report No.
    GAO/NSIAD-96-69BR). Washington, DC: Director, Military Operations and
    Capabilities, National Security and International Affairs Division B-270647.

Greenberg S (2004). Working through task-centered system design. In Diaper D &
    Stanton NA (eds.), *The Handbook of Task Analysis for Human-Computer Interaction*.
    New Jersey: Lawrence Erlbaum.

Hursh SR, Redmond DP, Johnson ML, Thorne DR, Belenky G, Balkin TJ, Storm WF,
    Miller JC, & Eddy DR (2004). Fatigue models for applied research in warfighting.
    Aviat Space Environ Med, *75*(3, Suppl.): A44–53.

Miller JC & Eddy DR (2009). *Operational Risk Management of Fatigue Effects II*.
    (AFRL Technical Report, Contract FA8650-06-C-6606). Brooks City-Base, TX:

Human Effectiveness Directorate, Biosciences and Protection Division, Biobehavioral Performance Branch.

Miller JC, Eddy DR, & Moise SL (2008). *24/7 Operational Effectiveness Toolset: Usability Assurance Plan*. (AFRL Technical Report, AFRL-RH-BR-TR-2008-0052, Contract FA8650-06-C-6606). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and Protection Division, Biobehavioral Performance Branch.

Miller JC, Eddy DR, Smith R, & Moise, SL (2009). *24/7 Operational Effectiveness Toolset: Shiftwork Scheduler Interface*. Technical Report, AFRL-RH-BR-TR-2009-0032, Contract FA8650-06-C-6606). Brooks City-Base, TX: Human Effectiveness Directorate, Biosciences and Protection Division, Biobehavioral Performance Branch.

Reason, J. (1990). *Human error*. New York: Cambridge

Shappell, SA & Wiegmann, DA (2000). *The Human Factors Analysis and Classification System–HFACS*. (DOT and FAA Technical Report No. DOT/FAA/AM-007). Oklahoma City, OK: FAA Civil Aeromedical Institute.

USAF Guide to Mishap Investigation (1987). AFP 127-1, Vol 1, 29 May 1987. No longer in print.

Other Documents Consulted

Aviation Safety Investigations and Reports (2004). (AF Manual No. AFMAN 91-223). Kirtland AFB, NM: HQ USAF/SE, available digitally on the AFDPO Website, http://www.e-publishing.af.mil.

Ground Safety Investigations and Reports (2004). AFMAN 91-224, HQ USAF/SE, 9700 G Avenue SE, Kirtland AFB NM 87117-5670, available digitally on the AFDPO WWW site at http://www.e-publishing.af.mil.

King R, Callister J, Patterson J, & Sipes W (1997). *Human Factors Handbook of Aircraft Mishap Investigation*. Human Engineering Division, Wright-Patterson AFB.

University Press.Safety Investigations and Reports (2006). AFI 91-204. HQ USAF/SE, 9700 G Avenue SE, Kirtland AFB NM 87117-5670, available digitally on the AFDPO WWW site at: http://www.e-publishing.af.mil.

APPENDIX A: NTI USER QUESTIONNAIRE

Years of active duty: _____   Years as mishap investigator: _____

Number of mishap boards or similar: _____   FAST user? Y   N

Rate the ease of application of the MIT to the intended task: the simplicity with which the MIT can be employed to determine whether fatigue was a factor in a mishap. In an ideal world, the interface would be totally natural and predictable in behavior. Nothing should obstruct your progress in completing this task.
1. Very acceptable (1)
2. Acceptable (2)
3. Borderline (3)
4. Unacceptable (4)
5. Very unacceptable (5)

Rate the performance of the MIT: the speed with which the interface responds to requests.
6. Very acceptable (1)
7. Acceptable (2)
8. Borderline (3)
9. Unacceptable (4)
10. Very unacceptable (5)

Rate the support information for the MIT: the information available to acquire, use and support the MIT. Encompasses initial instructions, user guides, tutorials, integrated assistance, context sensitive help.
11. Very acceptable (1)
12. Acceptable (2)
13. Borderline (3)
14. Unacceptable (4)
15. Very unacceptable (5)

Rate the MIT's function: the overall capabilities of the MIT.
16. Very acceptable (1)
17. Acceptable (2)
18. Borderline (3)
19. Unacceptable (4)
20. Very unacceptable (5)

Please discuss with the observer:
2. What were your objectives as you tested this interface?
3. Was the scope of the usability testing that you did adequate to meet your objectives?
4. Can you suggest another method of raw data entry that would reduce time, prevent entry errors, and provide greater awareness of data conflicts/errors?
5. Can you suggest other data editing methods that would work on a web page and would be more powerful for making changes?
6. Can you suggest other ways of displaying the mission log that would facilitate understanding, clarity, and accuracy of the situation?

7. Could the MIT report be formatted differently to better assist you in completing your mishap investigation and report?
8. Could the MIT graph be formatted differently to better assist you in completing your mishap investigation and report?

APPENDIX B: NTI F-PAS MIT USABILITY DATA COLLECTION FORM

Data:
1. On a separate screen, keep orderly, transcribable notes of the **pathways** the participants take, **problems** participants have and what participants **say** as they work. Definitions for the table, below:
2. *Number of subtask **assists***: When the participant cannot proceed on a subtask, the observer gives direct procedural help to allow the test to proceed.
3. *Number of subtask **errors***: Instances where test participant had to attempt portions of the task more than once.
4. *Number of subtask **reversals***: Number of times participant had to "back up" to find something on a previous screen that they needed on the current screen.
5. *Subtask **completion** (Y/N):* **Yes** = complete and correct achievement of subtask goal.
● *Problem **severity** (0/1/2):* **0** = no problem; **1** =minor (users are annoyed, but this does not keep them from completing the scenario);
   2 = show stopper (if we don't fix this, users will not be able to complete the scenario; and/or many users will be frustrated if we don't fix this; they may give up).

Start time:

Subtask	# Assists	# Errors	# Reversals	Severity	Completion
1. Enter user name and password at F/Pas site. Success = appearance of menu screen.				0 1 2	Yes No
2. Choose "Mishap Investigations" in "Custom Procedures" box. Success = appearance of MIT Intro screen.				0 1 2	Yes No
3. Choose "Download Mishap Investigation Worksheet." Success = appearance of worksheet PDF file. **Obs:** *Skip the Save/Open choice and give 4 printed copies to the participant.*				0 1 2	Yes No
4. Read the copilot scenario and make manual entries on the Mission Log Worksheets. Success = fundamental match with reference worksheet. *This is not software usability.*	NA	NA	NA	0 1 2	Yes No
5. Prepare to enter data. Success = return to MIT Intro screen.				0 1 2	Yes No
6. Choose "Do Analysis." Success = appearance of General Questions screen.				0 1 2	Yes No
7. Enter the AFSAS 6-digit number. Click "Next." Success = appearance of Investigator Name question.				0 1 2	Yes No
8. Enter the investigator's name. Click "Next." Success = appearance of role question.				0 1 2	Yes No
9. Select the participant role from the drop-down list. Success = match to worksheet.				0 1 2	Yes No
10. Click "Next." Success = appearance of location question.				0 1 2	Yes No
11. Select the participant location from the drop-down lists. Success = match to worksheet.				0 1 2	Yes No

Subtask	# Assists	# Errors	# Reversals	Severity	Completion
12. Click "Next." Success = appearance of Local vs. Zulu time question.				0 1 2	Yes No
13. Select Local or Zulu time base. Success = match to worksheet.				0 1 2	Yes No
14. Click "Next." Success = appearance of date question.				0 1 2	Yes No
15. Enter the mishap date. Success = match to worksheet.				0 1 2	Yes No
16. Click "Next." Success = appearance of time question.				0 1 2	Yes No
17. Enter the mishap time. Success = match to worksheet.				0 1 2	Yes No
18. Click "Next." Success = appearance of typical sleep time question.				0 1 2	Yes No
19. Check the "Change" box and click "Next." Success = appearance of typical bedtime screen.				0 1 2	Yes No
20. Change the typical bedtime to 22:00. Click "Next." Success = entry of 22:00 and appearance of typical wake time screen.				0 1 2	Yes No
21. Change the typical wake time to 06:00. Click "Next." Success = entry of 06:00 and appearance of Mission Log screen.				0 1 2	Yes No
22. Begin to enter data into the Mission Log. Success = noticing that this is the mishap day.				0 1 2	Yes No
23. Entry of data from Mission Log Worksheet for all days. *Obs: For each entry, there is a pop-up box containing relevant data. Make specific notes about **each activity, location and event** entry. **Summarize here.**				0 1 2	Yes No
24. Click on the "Mishap Analysis" button. Expect some questions about which button to click. Success = appearance of Mishap Analysis screen.				0 1 2	Yes No
25. Click to generate the report. Success = appearance of choice screen.				0 1 2	Yes No
26. Choose Local, Zulu or worksheet (default) time. ***Note:*** *presently, the report is generated only in Zulu time.*					
27. Click "OK" to generate the report. Success = appearance of report.				0 1 2	Yes No
28. Save the report to the Desktop. Success = appearance of file on Desktop.				0 1 2	Yes No
29. Return to the Mishap Analysis screen. Success = appearance of Mishap Analysis screen.				0 1 2	Yes No
30. Click to generate the graph. Success = appearance of graph.				0 1 2	Yes No
31. Save the graph to the Desktop. Success = appearance of file on Desktop.				0 1 2	Yes No

Subtask	# Assists	# Errors	# Reversals	Severity	Completion
32. Return to the Mishap Analysis screen. Success = appearance of Mishap Analysis screen.					
33. Save the file to the Desktop, naming it "Copilot-Ramstein." Success = appearance of file on Desktop.				0 1 2	Yes No
34. Save the same file to the Desktop as "Pilot-Ramstein." Success = appearance of file on Desktop.				0 1 2	Yes No
35. Return to Mishap Analysis screen by canceling or returning from the Windows save-file pop-up window. Success = appearance of Mishap Analysis screen.				0 1 2	Yes No
36. **Obs:** *Give 4 more printed copies of the Mission Log worksheet to the participant. Richard will add a button on this screen to allow the user to print more copies here.*				0 1 2	Yes No
37. Read the pilot scenario and make manual entries on the Mission Log Worksheets. Success = fundamental match with reference worksheet. *This is not software usability.*	NA	NA	NA	0 1 2	Yes No
38. Prepare to enter data by clicking "Questions" button. Success = appearance of Questions screen. *Note: the user will be at the end of the sequence of questions and will need to **back up** through the questions screens.*				0 1 2	Yes No
39. Do not change the AFSAS 6-digit number.				0 1 2	Yes No
40. Do not change the investigator's name.				0 1 2	Yes No
41. Select the participant role from the drop-down list. Success = match to worksheet.				0 1 2	Yes No
42. Do not change the participant location.				0 1 2	Yes No
43. Do not change the time base, the mishap date or the mishap time.				0 1 2	Yes No
44. Navigate to the typical sleep time question.				0 1 2	Yes No
45. Check the "Change" box and click "Next." Success = appearance of typical bedtime screen.				0 1 2	Yes No
46. Change the typical bedtime to 23:00. Click "Next." Success = entry of 23:00 and appearance of typical wake time screen.				0 1 2	Yes No
47. Change the typical wake time to 07:00. Click "Next." Success = entry of 07:00 and appearance of Mission Log screen.				0 1 2	Yes No
48. Begin to enter pilot sleep data into the Mission Log. Success = noticing that this is the mishap day.				0 1 2	Yes No

Subtask	# Assists	# Errors	# Reversals	Severity	Completion
49. Entry of pilot sleep data from Mission Log Worksheet for all days. For each entry, there is a pop-up box containing relevant data. Make specific notes about **each activity, location and event** entry. **Summarize here.**				0 1 2	Yes No
50. Click on the "Mishap Analysis" button. Expect some questions about which button to click. Success = appearance of Mishap Analysis screen.				0 1 2	Yes No
51. Click to generate the report. Success = appearance of report.				0 1 2	Yes No
52. Save the report to the Desktop. Success = appearance of file on Desktop.				0 1 2	Yes No
53. Return to the Mishap Analysis screen. Success = appearance of Mishap Analysis screen.				0 1 2	Yes No
54. Click to generate the graph. Success = appearance of graph.				0 1 2	Yes No
55. Save the graph to the Desktop. Success = appearance of file on Desktop.					
56. Save the pilot's file to the Desktop.				0 1 2	Yes No

End time:

APPENDIX C: MISHAP INVESTIGATION SCENARIO

Background:
This mishap occurred on 26APR1735Z landing at Ramstein with the co-pilot in command of the aircraft. The following activities were recorded for the two days prior and the day of the mishap.

1st duty period:
Show time was 3 hours before planned take-off time. The mission departed Travis at 23APR1648Z, arrived at Randolph at 23APR2101Z, and departed Randolph for Pope at 23Apr2350Z, landing at 24APR0220Z. The copilot reported taking a 45-minute nap (with a sleep quality of "poor") from 24APR0115Z to 0200Z on the flight from Randolph to Pope. (**Note:** he sleep quality function may not be available for the usability test. If not, just ignore sleep quality.)

The copilot was not required to attend the maintenance debrief on landing at Pope. He grabbed a quick meal and checked into the BOQ at 23APR2230L and was in bed falling asleep at 23APR2300L. (Pope is Z-4; ignore daylight savings time for this test.)

2nd duty period:
The copilot slept well (with a sleep quality of "good"), awakening at 24APR0830L. He cleaned up, went to breakfast at the club, and returned to his room to catch up on his professional military studies. He joined some friends stationed at Pope for lunch and then took in a matinee feature at the base theater. He came back to his room and checked out of the BOQ, grabbed a take-out meal, and reported to Ops at 24APR2100Z for pre-mission planning.

The mission departed Pope at 25APR0015Z and he remained awake during this leg. Mission landed at Lajes at 25APR0520Z and the crew cleared Ops at 25APR0600Z. The crew had breakfast together and then checked into the Lajes BOQ. He reported sleeping from 0600L-1030L (with a sleep quality of "fair"). (Lajes is Z-1.)

3rd duty period:
On awakening at 1030L, the copilot watched a local television variety show for about and hour and went to lunch at the club. He went to the gym in the early afternoon, followed by a nap in his BOQ room from 1600L-1700L (sleep quality "fair"). He watched television for a while and then joined the crew for dinner at the club at about 1900L. He went to bed about 2300L, falling asleep at about 2315L. He slept uninterrupted until awakening to his alarm at 0530L (sleep quality "good"), checking out of the BOQ at 0600L with the rest of the crew.

They had breakfast and reported to Ops at 26APR0915Z for the leg to Ramstein AB, Germany. The mission workload did not permit for napping en route. The mishap occurred at 26APR1735Z during landing at Ramstein, with the copilot making the landing.

After generating the report and viewing the graph for the copilot, save the copilot file. Then, click on 'File-Save As' and create a new file for the pilot. The mission times remain the same for the pilot, but the sleep times are different:

1st duty period:
The pilot did not nap on the flight from Randolph to Pope. He was was required to attend the maintenance debrief on landing at Pope. He had a quick meal, checked into the Pope BOQ and went to bed at 23APR2345L.

2nd duty period:
The pilot slept well (with a sleep quality of "good"), awakening at 24APR0730L. He did not nap during the leg from Pope to Lajes. He checked into the Lajes BOQ and slept from 0630L-1030L (with a sleep quality of "fair").

3rd duty period:
The pilot napped in his BOQ room from 1530L-1630L (sleep quality "fair"). Later, he went to bed at about 2330L and awoke at 0530L (sleep quality "good").

APPENDIX D: FULL FIRST INSPECTION EVALUATION

The SMEs specific usability comments and suggested modifications to the interface follow. We have listed them in the order of their occurrence for data entry in the MIT and used in the first 32 steps and the 39th step in the data sheet (Appendix B). Due to the time spent in analysis of the preceding steps, steps 33-56 (entry of second crewmember sleep data) were not performed. Steps without comments are included in this list for context.

- Enter user name and password at F-PAS web site.
- Choose "Mishap Investigations" in "Custom Procedures" box.
  - One SME was confused about the meaning of "Custom Uses" and why it was at the top of the items for him to choose. The special interfaces should be presented first or described better to make it easier for users to select the correct interface.
  - One SME was unsure what selection to make and was given assistance in selecting the correct button
- Choose "Download Mishap Investigation Worksheet."
  - One SME needed prompting as to what button to click.
- Read the scenario and make manual entries on the Mission Log Worksheets.
  - Add a "D - __ __" entry at the top of the form.
  - Add activity codes on form.
  - One SME was confused about what to do with show times vs. takeoff times. Were show times events or the beginning of a work period?
  - One SME missed the first show time. Show time is aircraft-type- and mishap-specific.
  - One SME did not enter work start and stop times, and was confused about what constituted a work period. Work should be considered an optional data entry capability because the model uses sleep times and does not need work time for computation.
  - One SME entered local times on the Zulu row and Zulu times on the Local row. He re-accomplished the worksheets.
  - After all participants completed the data entry on the printed form, the data entry forms were compared among the SMEs to ensure that all participants were working with a correct data entry sheet.
- Choose "Do Analysis."
- Enter the AFSAS 6-digit number.
- Enter the investigator's name.
- Select the participant role from the drop-down list.
- Select the participant location from the drop-down lists.
  - One SME had initial uncertainty about location entry (Cities, Bases, Lat/Lon).
- Select Local or Zulu time base.
  - The year should be included when displaying dates.
- Enter the mishap date.
- Enter the mishap time.
- Check the "Change" box.
- Change the typical bedtime to 22:00.
- Change the typical wake time to 06:00.

45

- Enter data into the Mission Log.
  - One SME failed to note that the initial page displayed was the mishap day. Perhaps add a transient screen that says "Going to the mishap day"? Or force a choice of which day to use for initial data entry?
  - One SME was annoyed by the incorrect display of the instructions because the screen resolution was set incorrectly.
  - SMEs did not know where to enter sleep/work data and needed instruction.
  - In entering the activity data, the SME desired to click and drag from right to left and the MIT display had some trouble staying with him and filling in the activity line correctly.
  - There was general uncertainty as to which activity slot to start/end on for sleep and work. This resulted in different participants getting different analysis results, by 15-minute increments, based upon how they entered the data.
  - For entering Event Type, the window did not show the time and the SME wanted to see it.
  - One SME was confused as to whether show time was an event.
  - One SME attempted to enter sleep as an event.
  - If an event location was entered first followed by its name, the location was replaced by "L/L" in the cell. The SMEs corrected errors made by the MIT. The location and event information should be entered in the same window.
  - A suggestion was made to remove the work interval from the work sheet since it was not needed by the model to project the performance effectiveness at the time of the mishap. However, it was pointed out that by knowing the work times, sleep times could be more easily constrained, thus reducing error.
  - The Location line was confusing to one SME. The Location line should be a display-only feature.
- Click on the "Mishap Analysis" button to see the analysis. Observations by SMEs:
  - Need the effectiveness circle in the dashboard.
  - Provide a way to place the dashboard in the system clipboard.
  - The dashboard was very intuitive and easily understood.
  - The "Chronic sleep debt" on the Dashboard will need to be changed to align with the AFSAS nanocodes.
- Click to generate the report.
- Choose Local, Zulu or worksheet (default) time.
- Click "OK" to generate the report.
  - The program crashed when the SMEs clicked on the "Report" button. After selecting "Back" the testing was continued.
  - Include AFSAS fatigue nanocodes (acquired from SMEs)
- Return to the Mishap Analysis screen.
- Click to generate the graph.
  - The graph needs a blue line to show sleep and a red line to show work.
- Return to the Mishap Analysis screen.
- Prepare to enter a different crewmember's data by clicking the "Questions" button.
  - The button labeled "Questions" was not understood intuitively by one SME. It was understood to be a help button.

A draft "Help" document was reviewed by the SMEs. Their comments included:
- Explain the Dashboard flags.
- Include AFSAS fatigue nanocodes (acquired from SMEs).
- For non-context-sensitive help, go to a table of contents (with anchors) at the top of the Help document.

The SMEs had some difficulty converting local time to Zulu, since they did not do it often enough to stay practiced. It was decided that prompts on how to find a Zulu calculator on the web should be made available to the investigators. If we gave them a button or URL in F-PAS, it would always have to be updated to make sure the site was always there. Also, there will be a general, tabular guide to time zones around the world printed on the reverse side of the Mission Log worksheet.

The general discussions among the observers and SMEs generated the following recommendations for the Mission Log:
- Make the Location and Activity rows display only
- Make all entries events, including sleep and work
- Add other event types such as show time
- Add an "Other" event type and specify if it has associated location or drug data
- Place tool tips on sleep/work intervals to show start, end and duration of the interval.
- Move the instruction pane lower and make it two lines, to fit at the bottom of the page.